

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/539,834  
Source: PCF  
Date Processed by STIC: 3/14/06

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/539,834

TIME: 11:06:02

Input Set : A:\10539834.txt

Output Set: N:\CRF4\03142006\J539834.raw

3 <110> APPLICANT: National Institute of Advanced Industrial Science and Technology  
 4 Hisashi NARIMATSU  
 5 Takashi KUDO  
 6 Akira TOGAYACHI  
 7 Toru HIRUMA  
 9 <120> TITLE OF INVENTION: GLYCOSYLTRANSFERASE, NUCLEIC ACID ENCODING THE  
 GLYCOSYLTRANSFERASE  
 10 AND METHOD OF TESTING CANCERATION USING THE NUCLEIC ACID  
 12 <130> FILE REFERENCE: 159-89 / YCT-902  
 14 <140> CURRENT APPLICATION NUMBER: US 10/539,834  
 C--> 15 <141> CURRENT FILING DATE: 2005-06-17  
 17 <150> PRIOR APPLICATION NUMBER: PCT/JP03/17030  
 18 <151> PRIOR FILING DATE: 2003-12-26  
 20 <150> PRIOR APPLICATION NUMBER: JP 380975/2002  
 21 <151> PRIOR FILING DATE: 2002-12-27  
 23 <160> NUMBER OF SEQ ID NOS: 20  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 1194  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Homo sapiens  
 30 <400> SEQUENCE: 1  
 31 atgcgctgcc ccaagtgcct tctctgcctg tcagcactgc tcacactcct gggcctcaaa 60  
 32 gtgtacatcg agtggacatc cgagtcccg gctcagcaagg cctacccag ccctcggggc 120  
 33 accccgccaa gccccacgcc agccaaccct gagccacccc tacctgccaa cctctccacc 180  
 34 cgcttggggc agactatccc gctgcccttt gcttactgga accagcagca gtggcggtcg 240  
 35 gggtccttgc ccagtgggga cagcactgaa acggggggct gccaggcttg gggggccgcc 300  
 36 gccgccaccg agatccctga cttegcctcc taccccaagg acctccgccg cttcttgcctg 360  
 37 tcagcagcct gccggagctt cccacagtgg ctgcctggag gtggtggcag ccaagtctcc 420  
 38 agctgctcag atactgatgt cccctacctg ctgttggccg tcaagtcaga accagggcgc 480  
 39 tttgcagaac gacaggccgt gagagagacg tggggcagtc cagctccagg gatccggctg 540  
 40 ctcttctcgc taggggtctcc ggtgggtgag gcggggcctg acctagactc actagtggcc 600  
 41 tgggagagcc gtcgctacag tgacctgctg ctctgggact tcctcgacgt cccattcaac 660  
 42 cagacgctca aagacctgct gctgctggcc tggctgggccc gccactgccc caccgtgagt 720  
 43 tttgtcttgc gagctcagga cgatgccttt gtacacaccc ctgccctgct ggctcacctg 780  
 44 cgggcccctg cacctgcctc ggcccgaagc ctctacctgg gtgaggcttt taccaggcc 840  
 45 atgcctctcc ggaagccagg aggacccttc tatgtgcccg agtccttctt cgaaggtggc 900  
 46 taccagcct atgcaagcgg ggggtggctac gtcattgccg ggcgcctggc accctggctg 960  
 47 ctgcgggcgg cagcccgtgt ggcacccttc ccctttgagg acgtctacac tggcctttgc 1020  
 48 atccgagccc tgggcctggt gccccaggcc caccagggct tcctcacagc ctggccagca 1080  
 49 gaccgcactg cggaccactg tgctttccgc aacctgtctg tggtaaggcc cctggggccc 1140  
 50 caggccagca ttcggctctg gaaacaactg caagacccaa ggctccagtg ctga 1194  
 52 <210> SEQ ID NO: 2  
 53 <211> LENGTH: 397  
 54 <212> TYPE: PRT

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/539,834

TIME: 11:06:03

Input Set : A:\10539834.txt

Output Set: N:\CRF4\03142006\J539834.raw

55 &lt;213&gt; ORGANISM: Homo sapiens

57 &lt;400&gt; SEQUENCE: 2

```

58 Met Arg Cys Pro Lys Cys Leu Leu Cys Leu Ser Ala Leu Leu Thr Leu
59 1          5          10          15
60 Leu Gly Leu Lys Val Tyr Ile Glu Trp Thr Ser Glu Ser Arg Leu Ser
61          20          25          30
62 Lys Ala Tyr Pro Ser Pro Arg Gly Thr Pro Pro Ser Pro Thr Pro Ala
63          35          40          45
64 Asn Pro Glu Pro Thr Leu Pro Ala Asn Leu Ser Thr Arg Leu Gly Gln
65          50          55          60
66 Thr Ile Pro Leu Pro Phe Ala Tyr Trp Asn Gln Gln Trp Arg Leu
67 65          70          75          80
68 Gly Ser Leu Pro Ser Gly Asp Ser Thr Glu Thr Gly Gly Cys Gln Ala
69          85          90          95
70 Trp Gly Ala Ala Ala Ala Thr Glu Ile Pro Asp Phe Ala Ser Tyr Pro
71          100         105         110
72 Lys Asp Leu Arg Arg Phe Leu Leu Ser Ala Ala Cys Arg Ser Phe Pro
73          115         120         125
74 Gln Trp Leu Pro Gly Gly Gly Gly Ser Gln Val Ser Ser Cys Ser Asp
75          130         135         140
76 Thr Asp Val Pro Tyr Leu Leu Leu Ala Val Lys Ser Glu Pro Gly Arg
77 145          150         155         160
78 Phe Ala Glu Arg Gln Ala Val Arg Glu Thr Trp Gly Ser Pro Ala Pro
79          165         170         175
80 Gly Ile Arg Leu Leu Phe Leu Leu Gly Ser Pro Val Gly Glu Ala Gly
81          180         185         190
82 Pro Asp Leu Asp Ser Leu Val Ala Trp Glu Ser Arg Arg Tyr Ser Asp
83          195         200         205
84 Leu Leu Leu Trp Asp Phe Leu Asp Val Pro Phe Asn Gln Thr Leu Lys
85          210         215         220
86 Asp Leu Leu Leu Leu Ala Trp Leu Gly Arg His Cys Pro Thr Val Ser
87 225          230         235         240
88 Phe Val Leu Arg Ala Gln Asp Asp Ala Phe Val His Thr Pro Ala Leu
89          245         250         255
90 Leu Ala His Leu Arg Ala Leu Pro Pro Ala Ser Ala Arg Ser Leu Tyr
91          260         265         270
92 Leu Gly Glu Val Phe Thr Gln Ala Met Pro Leu Arg Lys Pro Gly Gly
93          275         280         285
94 Pro Phe Tyr Val Pro Glu Ser Phe Phe Glu Gly Gly Tyr Pro Ala Tyr
95          290         295         300
96 Ala Ser Gly Gly Gly Tyr Val Ile Ala Gly Arg Leu Ala Pro Trp Leu
97 305          310         315         320
98 Leu Arg Ala Ala Ala Arg Val Ala Pro Phe Pro Phe Glu Asp Val Tyr
99          325         330         335
100 Thr Gly Leu Cys Ile Arg Ala Leu Gly Leu Val Pro Gln Ala His Pro
101          340         345         350
102 Gly Phe Leu Thr Ala Trp Pro Ala Asp Arg Thr Ala Asp His Cys Ala
103          355         360         365
104 Phe Arg Asn Leu Leu Leu Val Arg Pro Leu Gly Pro Gln Ala Ser Ile

```

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/539,834

TIME: 11:06:03

Input Set : A:\10539834.txt

Output Set: N:\CRF4\03142006\J539834.raw

```

105      370      375      380
106 Arg Leu Trp Lys Gln Leu Gln Asp Pro Arg Leu Gln Cys
107 385      390      395      397
110 <210> SEQ ID NO: 3
111 <211> LENGTH: 31
112 <212> TYPE: DNA
113 <213> ORGANISM: Artificial Sequence
115 <220> FEATURE:
116 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR
118 <400> SEQUENCE: 3
119 ctcaagctta tgcgctgccc caagtgcctt c 31
121 <210> SEQ ID NO: 4
122 <211> LENGTH: 31
123 <212> TYPE: DNA
124 <213> ORGANISM: Artificial Sequence
126 <220> FEATURE:
127 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR
129 <400> SEQUENCE: 4
130 ctgaattct cagcactgga gccttggggtc t 31
132 <210> SEQ ID NO: 5
133 <211> LENGTH: 20
134 <212> TYPE: DNA
135 <213> ORGANISM: Artificial Sequence
137 <220> FEATURE:
138 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for RT-PCR
140 <400> SEQUENCE: 5
141 gctgttggcc gtcaagtcag 20
143 <210> SEQ ID NO: 6
144 <211> LENGTH: 18
145 <212> TYPE: DNA
146 <213> ORGANISM: Artificial Sequence
148 <220> FEATURE:
149 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for RT-PCR
151 <400> SEQUENCE: 6
152 caggaagagc agccggat 18
154 <210> SEQ ID NO: 7
155 <211> LENGTH: 18
156 <212> TYPE: DNA
157 <213> ORGANISM: Artificial Sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Description of Artificial Sequence: probe for RT-PCR
162 <400> SEQUENCE: 7
163 cagaacgaca ggccgtga 18
165 <210> SEQ ID NO: 8
166 <211> LENGTH: 29
167 <212> TYPE: DNA
168 <213> ORGANISM: Artificial Sequence
170 <220> FEATURE:
171 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR

```

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/539,834

TIME: 11:06:03

Input Set : A:\10539834.txt

Output Set: N:\CRF4\03142006\J539834.raw

```

173 <400> SEQUENCE: 8
174 gccaaagctta catccgagtc ccggctcag                29
176 <210> SEQ ID NO: 9
177 <211> LENGTH: 29
178 <212> TYPE: DNA
179 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR
184 <400> SEQUENCE: 9
185 gccaaagctta aggcctaccc cagccctcg                29
187 <210> SEQ ID NO: 10
188 <211> LENGTH: 28
189 <212> TYPE: DNA
190 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR
195 <400> SEQUENCE: 10
196 cggaattctc agcactggag ccttgggt                28
198 <210> SEQ ID NO: 11
199 <211> LENGTH: 55
200 <212> TYPE: DNA
201 <213> ORGANISM: Artificial Sequence
203 <220> FEATURE:
204 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR
206 <400> SEQUENCE: 11
207 ggggacaagt ttgtacaaaa aagcaggctt ccccagccct cggggcaccc cgcca                55
209 <210> SEQ ID NO: 12
210 <211> LENGTH: 54
211 <212> TYPE: DNA
212 <213> ORGANISM: Artificial Sequence
214 <220> FEATURE:
215 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR
217 <400> SEQUENCE: 12
218 ggggaccact ttgtacaaga aagctgggtc tcagcactgg agccttgggt cttg                54
220 <210> SEQ ID NO: 13
221 <211> LENGTH: 29
222 <212> TYPE: DNA
223 <213> ORGANISM: Artificial Sequence
225 <220> FEATURE:
226 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR
228 <400> SEQUENCE: 13
229 gccaaagctta catccgagtc ccggctcag                29
231 <210> SEQ ID NO: 14
232 <211> LENGTH: 29
233 <212> TYPE: DNA
234 <213> ORGANISM: Artificial Sequence
236 <220> FEATURE:
237 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' primer for PCR
239 <400> SEQUENCE: 14

```

## RAW SEQUENCE LISTING

DATE: 03/14/2006

PATENT APPLICATION: US/10/539,834

TIME: 11:06:03

Input Set : A:\10539834.txt

Output Set: N:\CRF4\03142006\J539834.raw

```

240 gccaaagctta aggcctaccc cagccctcg                      29
242 <210> SEQ ID NO: 15
243 <211> LENGTH: 28
244 <212> TYPE: DNA
245 <213> ORGANISM: Artificial Sequence
247 <220> FEATURE:
248 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' primer for PCR
250 <400> SEQUENCE: 15
251 cggaattctc agcactggag ccttgggt                      28
254 <210> SEQ ID NO: 16
255 <211> LENGTH: 372
256 <212> TYPE: PRT
257 <213> ORGANISM: Homo sapiens
259 <400> SEQUENCE: 16
260 Thr Ser Glu Ser Arg Leu Ser Lys Ala Tyr Pro Ser Pro Arg Gly Thr
261 1 5 10 15
262 Pro Pro Ser Pro Thr Pro Ala Asn Pro Glu Pro Thr Leu Pro Ala Asn
263 20 25 30
264 Leu Ser Thr Arg Leu Gly Gln Thr Ile Pro Leu Pro Phe Ala Tyr Trp
265 35 40 45
266 Asn Gln Gln Gln Trp Arg Leu Gly Ser Leu Pro Ser Gly Asp Ser Thr
267 50 55 60
268 Glu Thr Gly Gly Cys Gln Ala Trp Gly Ala Ala Ala Thr Glu Ile
269 65 70 75 80
270 Pro Asp Phe Ala Ser Tyr Pro Lys Asp Leu Arg Arg Phe Leu Leu Ser
271 85 90 95
272 Ala Ala Cys Arg Ser Phe Pro Gln Trp Leu Pro Gly Gly Gly Ser
273 100 105 110
274 Gln Val Ser Ser Cys Ser Asp Thr Asp Val Pro Tyr Leu Leu Ala
275 115 120 125
276 Val Lys Ser Glu Pro Gly Arg Phe Ala Glu Arg Gln Ala Val Arg Glu
277 130 135 140
278 Thr Trp Gly Ser Pro Ala Pro Gly Ile Arg Leu Leu Phe Leu Leu Gly
279 145 150 155 160
280 Ser Pro Val Gly Glu Ala Gly Pro Asp Leu Asp Ser Leu Val Ala Trp
281 165 170 175
282 Glu Ser Arg Arg Tyr Ser Asp Leu Leu Leu Trp Asp Phe Leu Asp Val
283 180 185 190
284 Pro Phe Asn Gln Thr Leu Lys Asp Leu Leu Leu Leu Ala Trp Leu Gly
285 195 200 205
286 Arg His Cys Pro Thr Val Ser Phe Val Leu Arg Ala Gln Asp Asp Ala
287 210 215 220
288 Phe Val His Thr Pro Ala Leu Leu Ala His Leu Arg Ala Leu Pro Pro
289 225 230 235 240
290 Ala Ser Ala Arg Ser Leu Tyr Leu Gly Glu Val Phe Thr Gln Ala Met
291 245 250 255
292 Pro Leu Arg Lys Pro Gly Gly Pro Phe Tyr Val Pro Glu Ser Phe Phe
293 260 265 270
294 Glu Gly Gly Tyr Pro Ala Tyr Ala Ser Gly Gly Gly Tyr Val Ile Ala

```

VERIFICATION SUMMARY

DATE: 03/14/2006

PATENT APPLICATION: US/10/539,834

TIME: 11:06:04

Input Set : A:\10539834.txt

Output Set: N:\CRF4\03142006\J539834.raw

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date